DOCUMENT RESUME

ED 454 861 IR 058 152

AUTHOR Beacom, Matthew
TITLE Crossing a Digital Divide: AACR2 and Unaddressed Problems of Networked Resources.
PUB DATE 2000-11-00
NOTE 22p.; In: Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web (Washington, DC, November 15-17, 2000); see IR 058 144.
AVAILABLE FROM For full text: http://lcweb.loc.gov/catdir/bibcontrol/beacom_paper.html.
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Access to Information; *Cataloging; Change; *Information Utilization; *Internet; *Knowledge Representation; Standards
IDENTIFIERS *Anglo American Cataloging Rules 2; *Electronic Resources

ABSTRACT

This paper discusses the following four changes in how we use sources of information or representations of knowledge: (1) from tangible to intangible media--how sources of information and representations of knowledge are manifested or packaged on the Internet; (2) from books and journals to services and databases--how such knowledge packages are published and distributed on the Internet; (3) from buying to leasing--how access to knowledge packages is controlled on the Internet; and (4) from ascertaining to using, a new purpose for the catalog--how we help others to use knowledge packages on the Internet. The magnitude of these changes is evaluated, some connections to AACR2 (Anglo American Cataloging Rules, 2nd edition) are detailed, and 12 changes to AACR2 are suggested. (Contains 22 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 454 861

Crossing a Digital Divide: AACR2 and Unaddressed Problems of Networked Resources

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
B. Wiggins
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Matthew Beacom

A paper for the conference "Bibliographic Control for the New Millennium" held in Washington, DC at the Library of Congress, November 2000.

Final version

INTRODUCTION

The advent of the World Wide Web marks profound changes in how we use sources of information such as databases, indexes, and archives and how we use representations of knowledge such as maps, pictures, sounds, books, and journal articles. We are using the Web to change how we communicate with one another-how we read and write, how we speak and listen. We are using the Web to change how we do business-how we make and assemble things, how we buy and sell things. We are using the Web to change how we entertain ourselves. We are using the Web to change how we work. In short, we are using the Web to change our culture, to change how we live.

For librarians and those whom we serve, the most important changes in technology and society may be in how such sources of information or representations of knowledge are created and used. Specifically, these are changes in how knowledge is packaged or represented in an enduring physical form, how those physical forms or packages of knowledge are published and distributed, how their use is controlled or restricted, and how these packages of knowledge can be found, used, and saved for later use. These changes are profoundly affecting catalogers, catalogs, and catalog users.

A few of the chief topics of concern include "content versus carrier," "multiple versions," the purposes of the catalog, and its role in a networked information environment. By "content versus carrier" I refer to those problems or issues that affect how we relate the intellectual or artistic content of a representation of knowledge to the carrier or physical format that embodies it when there is a one to one relation between content and carrier. By "multiple versions" I refer to same issues of that relation between content and carrier when the relation between content and carrier is one to many.

In response to the Web and the cultural changes associated with it, librarians and their peers are thinking

IR058152

BEST COPY AVAILABLE

anew about how we enable persons to have access to sources of information and representations of knowledge. To that end, we are re-examining cataloging, catalogs, and our own role as intermediaries between those objects that embody information or knowledge-what Arlene Taylor calls "information packages" in her 1999 book, *The Organization of Information*-and those persons who would use them, our patrons, readers, or users (p. ??).

Librarians and their kin have been actively responding to the Web. Their responses include the following. MARBI changed MARC to accommodate access to remote electronic resources. IFLA published *Functional Requirements for Bibliographic RecordsAACR2* chapter 9 (on computer files) with the *ISBD(ER)*. The Joint Steering Committee (JSC) sponsored the International Conference on the Principles and Future of AACR in 1997 to explore how we might fundamentally revise the cataloging rules. Among the follow-up projects from that conference are a logical analysis of *AACR2* by Tom Delsey; a set of recommendations relating to issues of seriality in AACR by Jeans Hirons, et al.; and revisions to *AACR2* rule 0.24 that may establish the primacy of content over format for catalogers and user by CC:DA and the JSC. In light of all this thought and action, what problems of networked resources and *AACR2* could possibly be called unaddressed? The issues I discuss here, thus, are not literally unaddressed. For many others have considered, written, and acted on them. We are not, however, finished thinking about and acting on these issues.

In this paper, I discuss four changes in how we use sources of information or representations of knowledge, briefly evaluate the magnitude of these changes, detail some connections to *AACR2*, and suggest a few changes to *AACR2*. I address each of these four changes from the perspective of one who asks what is the relation between these changes and *AACR2*, between these changes and how libraries and librarians enable others to gain access to information and knowledge?

The changes

- From tangible to intangible media: how sources of information and representations of knowledge are manifested or packaged on the Internet
- From books and journals to services and databases: how such knowledge packages are published and distributed on the Internet
- From buying to leasing: how access to knowledge packages is controlled on the Internet
- From ascertaining to using, a new purpose for the catalog: how we help others to use knowledge packages on the Internet

Following the discussion of these four changes, I make 12 recommendations for changes to *AACR2*.

The size of the changes

Before discussing these changes individually, let's turn to estimating the magnitude of these changes. How big are these changes? Do they matter a little, a lot, or do they change everything? Christine Borgman, in her fine book, *From Gutenberg to the Global Information Infrastructure*, wisely

distinguishes continuous or evolutionary patterns of change from discontinuous or revolutionary patterns of change. Her analysis of the technological and social change we are living through leads her to a reasonable view that she calls "co-evolutionary." (p. ??) Her view emphasizes the mixed and dynamic nature of the various responses of people and the organizations to technological change and the unanticipated consequences of such responses on further technological and social change. So how big are these changes? Overall, I argue they are radical changes that will in time transform how we create and use knowledge packages. To play on Borgman's vocabulary, let's call it "co-revolutionary." It is not an absolute break with the past, but it is changing everything. We are crossing a digital divide.

THE CHANGE FROM TANGIBLE TO INTANGIBLE MEDIA

On one side of this digital divide we have such traditional media as books and journals for texts, photographs and films for images, records and tapes for recorded sounds. All are *tangible* objects that contain or carry information or knowledge-the intellectual or artistic content. As librarians we are very familiar with this. Catalogers who work with knowledge packages that do not fall into simple types of materials know this junction of content and format as the content versus carrier problem. The things we catalog, what I'm calling knowledge packages-books, online databases, maps, e-journals, recorded music, digital archives, journals, corporate Web sites, etc.-are complex objects. They are mixtures of information or knowledge and physical format or carrier. Until a short time ago, everything we knew about knowledge packages was based on our experiences with traditional-*tangible*--media.

On the other side-the one we are crossing to-we have the new media, the Web, the Internet. The new media are *intangible*, untouchable. We still have such familiar kinds of content as texts, images, and sounds, but with a difference. Our experience of the texts, images, sounds, etc. carried by the new media are at one remove from ourselves-mediated by our computers. We can do wonderful things with the new media, but at a price. We can no longer touch them. We based *AACR2* on the idea of cataloging the item in hand. And now the things-knowledge packages-we catalog can't be handled.

Yet things on the Web are still things, after all. The things we catalog need not be tangible to be *things*. The documentalists, who flourished earlier in this century, understood this. Michael Buckland in his 1991 article, "Information as Thing" quotes, Briet's definition of a "document" as any concrete or symbolic indication of a physical or mental phenomenon that has been recorded for reconstructing or recreating that phenomenon." (p. 355) Although similar in vocabulary to the Internet era idea of document-like objects, it is actually the reverse. Contrary to what their name implies, documentalists were expanding the idea of document to include all representations of knowledge within a comprehensive concept. The idea of document-like objects as applied in the Dublin Core initiative restricts its scope to those representations of knowledge that have the qualities or characteristics of documents.

Knowledge packages on the Web do not lack physical qualities. They lack some familiar qualities and have some we are not used to. Web things--changeable, ephemeral, and adaptable--can be lost, found, or hidden, bought, sold, or leased, corrupted, destroyed, or preserved, known, cited, or used. These Internet-

based knowledge packages retain many recognizable characteristics. For example, they are still intellectually recognizable as particular types of knowledge packages such as reports or poems, drawings or music, census data sets or road maps. The move from touchable to untouchable media fundamentally alters how we act on and use knowledge packages. *AACR2* defines some rules on one particular use of knowledge packages-the creation of metadata surrogates or catalog records for use in library catalogs.

The change to how representations of knowledge are packaged-the change from tangible to intangible media--encounters *AACR2* most immediately in rule 0.24, but also in the physical description area (ISBD area 5), and in the materials specific details area of *AACR2*. How networked resources are packaged profoundly affects how we understand and resolve the relation between content and carrier. The potential for multiple packages of the same or nearly the same content in networked resources greatly increases the scale of multiple version problems. These problems are central to bibliographic control of networked resources.

A cardinal principle

Let's look at *AACR2* rule 0.24 in more detail. Before this digital divide, we based description on the "physical" form of the item in hand. But by physical form we meant the categories referred to by chapters 2 to 12 of *AACR2*, the types of material. Despite the mixing of intellectual and physical formats, this arrangement worked well for the most part. Important exceptions include serials and reproduction microforms. The issues or problems with serials and reproduction microforms are well known. The treatment of seriality in *AACR2* is being significantly but not *radically* revised now. Reproduction microforms and the larger issues of reproductions and multiple versions in the digital environment have yet to be successfully addressed. The discussions now underway in the cataloging community including a draft LCRI 1.11A Non-Microfilm and Electronic Reproduction is unsuccessful primarily because of our failure to define what an electronic reproduction is. Because of the ease with which networked resources can be re-purposed, multiple version issues regarding bibliographic control and user services are approaching crises.

Content and carrier

Rule 0.24 has been revised now so that catalogers are directed to bring out all the characteristics of the item being described (i.e. formats detailed in *AACR2* chapters 2-12.) This is a solid improvement. It represents a big response to the change from tangible to intangible resources, but it is not the radical change that we will soon need.

There is an element of abstraction to all things. The saying "the eyes see what the mind knows" is a testament to the mixing of physical and mental in what we call things. A book is an *idea* as much as it is a tangible object. A Web site or e-book is also an idea as much as it is an intangible object. The abstraction level is just a little higher for us because we can't touch an e-book or a Web-site. Current practice is mixed for both analog and digital media. 0.24 classically says to base the description on the item in hand, in effect catalog the manifestation. But our practice is mixed. For example, the Library of

Congress's microform practice vs. *AACR2*-conformant practices such as that at the National Library of Canada. Or, the CONSER single record practice-a form of dashed-on entry for the 21st century-and the separate record practice.

Content and carrier cannot be separated without breaking the link between the work and the item. That link is vital to successfully serving the user who needs to have some particular item that is the right work, the right expression, and the right manifestation. If we can't break the link between the work and the item without harming the user, what do we do?

The distinction that we need to make is between analog and digital, stand alone and networked, tangible and intangible. By using the types of material defined in chapters 2-12 of *AACR2* as primary types we continue to confuse types of carriers with types of content and with modes of publication over time. Delsey's suggestion that *AACR2* be reorganized by ISBD area is a powerful idea. The test case reorganization done by Library of Congress's Cataloging Distribution Service unit shows that this reorganization is not trivial or mechanical. Serious intellectual effort to reorganize *AACR2* by ISBD (or ISBD-like) area is needed now. This approach may also be described as creating a super chapter 1 of *AACR2*. Practicing catalogers need guidelines for cataloging traditional types of materials or other categories of knowledge packages. These guidelines must be based on the rules, but need not be part of the rules themselves. In other words, instructions for cataloging some of the materials now in *AACR2* chapters 2 to 12 may best be separately published as material specific guidelines and not rules.

Multiple versions

How do we address multiple versions or multiple formats at the level of the records or surrogates that we create for our catalogs? For the multiple version issues, the big question is what things should our records be surrogates for? Should we make surrogates or records for the content itself abstracted from its carrier or should we make surrogates for the content *and* its carrier? In the terms of the *Functional Requirements for Bibliographic Records*, the question is should we catalog each expression of a work or each manifestation of the work? In current practice and rules, we mix content and carrier in flat surrogates or records. In the terms used in Functional Requirements for Bibliographic Records, each catalog record we make for a given knowledge package generally mixes information about the package at four levels: the work, the expression, the manifestation, and the item. Imagine a text in 6 formats: XML, SGML, pdf, postscript, on a DVD, and in print. Should that be described in 1 surrogate, 2, 3, or 6? If 1, how are the manifestations articulated. If more than 1, how is the principle of division made clear? If 6, how are the bibliographic relations among the 6 manifestations described? Further, by what criteria will we decide? If the content is the same, shouldn't it all be on one record? That is, in effect, the notion of cataloging the expression rather than each manifestation. But is content the only relevant criterion?

The chief contra-argument to cataloging the expression is that content is *inescapably* joined to carrier. This union of content and carrier makes the knowledge packages that our users actually use. The combination of content and carrier is what makes knowledge *usable*. These knowledge packages are the objects around which libraries-servers, buildings, networks, staffs, services, collections, and purposes-are

built. Libraries collect things, knowledge packages that mix item, manifestation, expressions, and work aspects into a whole. To meet its purposes, the library catalog needs to make those things the library has accessible to users. It does not matter whether they are on a shelf or on a server. The focus of on the thing itself is still vital. *Representations of knowledge are things*. When we lose sight of this fact, we lose our way.

Records for knowledge packages in multiple formats

Some catalogers want to split up knowledge things-describing each manifestation separately. Others want to lump knowledge things together-describing all manifestations of an expression together. Each group thinks its way is best, especially best for users. But the dichotomy of splitters and lumpers is a false one. It is not a matter of either/or. It is a matter of when and where. The criterion is what can we do that serves the user best.

The classic example of this split *and* lump approach is described in the *Guidelines for Bibliographic Description of Reproductions*. A tiered record approach both splits and lumps. One tier describes the expression (and provides work level access points), another tier describes the manifestations (and provides manifestation level access points.) This powerful model deserves renewed interest and effort. This model would solve our multiple version problems. However, it is not without problems. Among the best known are the need for compatibility with older records, with MARC formats, and with legacy OPAC and bibliographic utility systems. The need for backward compatibility may be one of the strongest pulls on librarians to adapt *AACR2* rather to jump to a born digital metadata scheme like the Dublin Core. We have an installed user base that we don't want to and can't abandon, and we have institutional commitments to servicing analog materials.

There are other ways to split and lump. We could split at the point of record creation and lump at the point of display. Our rules could dictate that we split at the record level by cataloging each manifestation, and lump at the display level by linking each manifestation record into a integrated display of expression level and manifestation level information. Of course, this approach has its own disadvantages. Our OPACs would have to intelligibly and flexibly show bibliographic relationships among records and link (conceptually and mechanically) across records. Our records would need subtle and robust areas for managing relationships and linkages. The Web just happens to be a really suitable environment for doing both of these things. It may be far more possible to do this in the next 10 years than it was in the last 10.

CC:DA's recent recommendation on 0.24 to the Joint Steering Committee (JSC) takes another approach. Split sometimes and lump other times; make the choice based on a list of major/minor changes that are to be appended to *AACR2*. The most recent CC:DA recommendation to the JSC builds on the earlier recommendation by devising a list of major/minor changes that would guide catalogers in deciding when to create new records. This approach will not work. The list of changes is not the tool catalogers need. The tool we need is a coherently conceived record structure, such as the tiered or linked record structures mentioned above. The question cataloger's need to ask is not-- when do I make a new record? The question to ask is how do we effectively distinguish and display work, expression, manifestation, and

item level information to users.

THE CHANGE FROM BOOKS AND JOURNALS TO SERVICES AND DATABASES

Changes to how packages of knowledge are published and distributed encounter *AACR2* most obviously in the publication, distribution, etc. area (ISBD area 4). But the impact is not limited to this area. Indeed, the impact of these changes is far broader than conventions for recording the places, names, and dates associated with publishing.

I address six aspects of this change and its affect on cataloging.

- What does "published" mean on the Internet?
- What are the consequences of the Internet flood of information sources?
- How does reference linking change cataloging?
- What new means of publishing and distributing representations of knowledge are likely to dominate the Web?
- What is the future for books in a media environment so conducive to interactive multimedia and continuous updating?
- What is value to users of imprint information (ISBD area 4) in the digital era?

What does "published" mean on the Internet?

Changes in what it means to publish affect *AACR2* fundamentally. The change from traditional media to digital networked media disrupts our understanding about what is and is not published, about what it means to publish. The dictionary says publish means "to make generally known," "to place before the public." This is fundamental, but only part of what it means to publish.

Our understanding of publishing is complex. Publishing is an intellectual, social, economic, and technological phenomenon. Our understanding is tied to the central distinction between public and private spheres of life and blurred by phenomena such as gray literature and invisible colleges. The Internet and our particular uses of it are driving changes in the public and private distinction, greatly increasing the visibility of gray literature, and through peer to peer networking making invisible colleges on a global scale possible.

Our understanding of publishing is more fully developed through the concepts expressed in words like original, copy, edition, impression, and reproduction. These concepts are traditionally associated and frequently used in our work. These familiar and traditional concepts have been built on our experience of analog formats. As we cross the digital divide, we extend these concepts to digital, networked knowledge packages or things in order to keep control of the new materials. We make do, innovate, and adapt.

In some library and publishing ventures these adapted working definitions have been useful. The re-publications of journals and books by JSTOR and netLibrary are examples of publishers extending the

use of these familiar publishing concepts to digital networked materials. Other ventures such as pre-print databases, e-journal aggregations, and personalized services that may replace traditional textbooks such as those offered by MetaText suggest new modes of packaging and distributing recorded information and knowledge.

Will these extensions or adaptations be for naught? In the context of the Internet, applying such familiar concepts of "original," "copy," "edition," "impression," and "reproduction" is often of doubtful value. For example, in the analog age the number of copies of a knowledge package made for distribution is limited. In the digital age, copy is more likely to be a verb than a noun. This is a small but telling difference. In the analog age, these terms represent fairly precise concepts; in the digital age, they become metaphors, new parts of speech, or anachronisms.

Since the Internet makes it so easy "to place things before the public," some have argued that we should treat everything on the Internet as published. However, the published or not published division can be made and may need to be. *Digital Dilemma: Intellectual Property in the Information Age* raises these issues clearly and makes a strong case that on this side of the digital divide publishing will still be a complex and nuanced phenomenon. We may include on the not published side such things as author's drafts, notes and other materials not used in finished products, as well as "private" material like e-mail or calendars or digital diaries. In revising *AACR2*, we need to decide how we will make and use distinctions between published and unpublished digital networked materials we may add to our collections and catalog.

For example, we may wish we could treat all online manifestations of some content as reproductions of an analog original form and use LC's microform practice to guide our cataloging. On the other side of the digital divide, print is dethroned. Print becomes just another output option, one that can be invoked or not invoked by the publisher, a wholesale reseller, a retailer, a library, or a reader. If there is an "original," it is online. We need to develop a new vocabulary and new concepts out of our analog *and* our digital experiences. And we need to use these ideas in *AACR2*. Work, expression, manifestation, and item have already been mentioned. These terms and ideas take us a long way. Problems with our concept of reproductions have also been mentioned and it requires further work.

The Internet flood of information sources?

A more pressing consequence of the increased ease of publishing is the sheer volume of materials on the Web. We are experiencing, in part, a tidal wave of gray literature. (Another portion of the wave is the result of a global village effect, e.g. every town's newspaper is online and available at any computer.) On the Internet, the distinction between published materials and gray literature is weakened. So much gray matter is so easy to find on the Web that more formally published material is lost or obscured. (Much of that formally published material is also hidden behind access restricting checks.) Metadata developments like the Dublin Core are partially predicated on this blurred distinction. Too many people are making too much material public through too many channels or outlets for traditional methods of bibliographic control such as library catalogs or national bibliographies to suffice. In revising *AACR2*, we need to

decide what relation library catalogs will have to the Internet. I address this aspect more fully below in the section on the changing role of the catalog, but the key word is *selection*.

The catalog and reference linking

Content can be re-packaged and leased many times to many such groups because the Internet makes it easy for many different agencies to license the same content to many different groups. This is one source of the proliferation of knowledge things and the records that describe them. Clearly, this can be a collection development issue for libraries-how may times do you want to buy the same content for your user group? But for the makers and users of catalogs, a defining aspect of networked resources adds a twist to the multiple version issues. URLs are not universal. The URL that links the resource described in one record only works for members of the licensed user group. This is no surprise to many and is one reason why MARBI defined the electronic location field (MARC tag 856) in both the bibliographic and the holdings formats.

MARBI has addressed the need for records or surrogates of online resources to link directly to the resources themselves. *AACR2* has not. URLs are often consider something like a shelf location or other completely local information and thus outside of any concern by *AACR2*. Without creating rules for making hyperlinks from surrogates to the resources themselves, *AACR2* breaks a linkage that is a defining characteristic of networked information. Where should this kind of information be: in notes, in standard numbers, a new section, or in a general rule and then added throughout the code as needed? I recommend the later. *AACR2* must explicitly address hyperlinks such as URLs, URNs, and others in the cataloging rules. ISBD needs to address this too. Otherwise in a digital world, the makers and users of digital media will ignore *AACR2* (and the ISBDs). Dublin Core is designed for Web-based knowledge packages. MARC has adjusted to the Web with the 856 field. The makers and user of *AACR2* (and the ISBDs) must recognize the critical importance of reference linking in a networked information environment. Redesign the rules to fit a publishing environment of pervasively inter-linked knowledge packages and a metadata environment of similarly inter-linked surrogates or records.

New publishing and distributing methods

Networked resources may be packaged in new ways. New bibliographic entities and new bibliographic relationships are native to the digital, networked environment. E-journal aggregations are one example. Are they convenient bundles of journals or are they precursors to new delivery mechanisms for articles? Article databases are replacing journals as the dominant *delivery mechanism* for articles. Journals are not likely to just go away. Their roles will change. They will continue to have powerful editorial functions with resulting value as brands and as a useful search limit term. Their function as devices for article delivery to the user will lessen in importance. Such a transformation in publication practices would significantly affect *AACR2* chapter 12 both in its current form (i.e. Serials) and in its emerging form (i.e. Continuing Resources.)

In general, the impact of new kinds of knowledge things on the Internet on *AACR2* is to undermine

AACR2's extensible structure. Although *AACR2* is designed to adapt to new formats of materials by adding chapters, adding new chapters for proliferating e-formats (tangible and intangible) is not a viable choice. Second, we have mixed up physical formats with characteristics and qualities. For example, seriality is a condition not a format. It is potentially applicable to any knowledge thing we can imagine: texts, images, cartographic information, sound recordings, etc. The current recommendations before the Joint Steering Committee now recognize this, but for many reasons, mostly practical ones, the changes are mainly contained within the chapter for the serials format, chapter 12. We are still trying to compartmentalize seriality, to treat it as a format not as a range of conditions or characteristics that might apply to any knowledge package.

Furthermore, digital networked resources are at least as likely to be *blends* of what we have traditionally called formats as they are likely to be, shall we say, single malts. Now we use adjuncts to *AACR2* like *Guidelines for bibliographic description of interactive multimedia and Cataloging Internet Resources: a Manual and Practical Guide* to retrofit the rules. We have made modest changes to the rules as in the revision of *ISBD(CF)* to *ISBD(ER)* and the recent efforts to harmonize *AACR2* chapter 9 with *ISBD(ER)*. Now, on this side of the digital divide, we need to rethink our rules with the networked environment as the technological and social base for communication. To accomplish this a thoroughgoing revision of *AACR2*, such as that suggested by Delsey is needed. We are making progress. We are moving quickly for our profession, but slowly for the larger networked environment in which we now find ourselves.

One specific impact may be seen in possible responses to the development of article databases. Two options come to mind. Return to article level cataloging. This is possible but unlikely to be a successful strategy. Our experience in the past century with third party journal article indexing has demonstrated its efficacy relative to cataloging journals article by article. A second option is to link cataloging and indexing information in ways that the user sees as seamless. Developments with reference linking tools like jake and SFX indicates the power of such a smart, scalable approach.

For *AACR2* to support such deep or integrated linking, the cataloging community needs to add a new area on linkage and relationships to *AACR2*. In *AACR2* hypertext or hyperlinks-and the technological and social environment that supports it and expects it-does not exist (except perhaps in chapter 9 and only grudgingly and implicitly.) It should be a fundamental principle of cataloging in a digital age that all records and other metadata surrogates should be designed to link to other surrogates that describe the same resource at different levels of granularity or other related resources (even those that use other metadata schemas. Such linkages can be applied to analog and digital materials.

New resources, new qualities

Networked resources have new qualities. For example, digital networked resources are egregiously updateable. In *AACR2*, updateable publications such as looseleafs are marginal at best. For networked resources, updating publications may become the dominant pattern. Such resources are also open to combinations of multiple media (text, images, sound, etc.) in one publication in ways that are unthinkable in print or other analog formats. Furthermore, such resources are inherently linkable. One

example of this shift may be a move from monographs to interactive services. This change may be indicated by NetLibrary's development of its MetaText product. This product is a set of Web-based communication and analysis tools with interactive multimedia content-enriched textbooks as its content. (NetLibrary)

The development and dominance of services over distinct objects may lead us beyond what a catalog can contain. Or perhaps it is only an issue of granularity. We can catalog the services as entities and not the shape-shifting products one can produce on demand from such services. This is similar to collection level cataloging. But how do we contribute to making materials below the level of the whole service accessible? This is a new responsibility for the catalog and for *AACR2*. The key is cross-profession collaboration and inter-linked metadata standards. We are doing that now with archival finding aids that use the EAD DTD. The archival collection can be cataloged using *AACR2* and expressed in MARC for transport and use in OPACs. The record includes a hyperlink to the EAD- encoded finding aid for that collection. Librarians use one standard for the catalog record and archivists use another for the finding aid. Users benefit from metadata created by two different but related support communities. The assumption of catalogers must be that the surrogates or records they create will be used in conjunction with other forms of metadata. *AACR2* revisions must explicitly declare this assumption, and it must design rules around its consequences.

The value of imprint information

In a digital networked environment what is the value to users of the information recorded in the publishing, distribution, etc. area? Does the Internet affect the value of the imprint information and its use? For example, does *place* of publication matter online? The place of publication may not matter at all, or it may matter in new ways. Users may find new values or new uses for imprint information. If the former, why record it? If the latter, will those new uses change how and what we will record? Imprint information may become more important for access than for description and identification. Citation practices may change from the conventions developed in an analog age. (National bibliographies may also change their practices as the Internet enables increasing globalization of enterprises like commercial publishing.) *AACR2* needs to address this particular question but also more generally ask what is the role of *transcription* in an era of networked resources. Transcription has never been and should not be an end in itself. It has always served the function of identification by enabling the surrogate to mirror the resource it describes.

This shift in the value of imprint information isn't just a matter of digital form. It is also a consequence of globalization. The named publisher may be little more than a brand name in a multinational media conglomerate, and places of publication have been proliferating for print publications, too. But there is an analog vs. digital divide here, too. In the analog era, it is used in part as evidence to identify the manifestation and the expression-the edition. In digital era, are such indirect indications of edition as useful? The *name* the publisher may matter to the user and the library less than the name of the e-seller, the e-aggregator or e-jobber? Electronic materials may be re-packaged by so many vendors that the publisher may not matter to users or the library as much as the vendor may. The agency with which the library or the user has signed a licensing agreement may be far more important for description and access

in the digital era than the publisher.

THE CHANGE FROM BUYING TO LEASING

Changes to how one controls access to packages of knowledge currently encounters *AACR2* primarily in the note area (ISBD area 7). But its impact is not limited to this area. The change from an environment dominated by buying and selling knowledge things and controlled by copyright law to one dominated by leasing and controlled by licensing agreements is fundamental. It alters the relations between the library and the things it collects, between the library and its users, and between the knowledge things themselves and their use and usability. In the analog era, communication and scholarship ranged across a fair use commons. On this side of the digital divide, the fair use commons is being claimed and fenced in. The impact of these changes on *AACR2* is critical to the relevance of the rules within an environment where licensing agreements control the exchange of information.

Access restrictions are uncommon with many analog materials. A notable exception is archives of unpublished materials. In the analog era, copyright is the defining rights management paradigm for the relations between publishers and users. Habits of sale and use have developed in this relatively stable technological, social, and legal context. Copyright itself is an extrinsic context within with items are bought and sold. With networked resources, though, access restrictions are not only commonplace, but also vital characteristics of digital objects. Although they are not strictly speaking intrinsic qualities, access restrictions or, more generally speaking, rights management conditions are profoundly in-twinned with the *use* of digital networked resources.

Notes about access restrictions are helpful to users. Users might not read them, but they are better than nothing is. A combination of universal note and local note is often most useful. Without them the user has no hope of knowing what items in any search result are or are not accessible to them until *after* they attempt to retrieve each item. An access forbidden message will let them know at some point, say when they try to see the full text of a particular article. However, notes alone are inadequate to rights management in an environment dominated by leasing.

AACR2 has no area for dealing with rights management. The notes area and terms of availability section are inadequate substitutes for a rights management area. Access restrictions and rights management must be explicitly addressed by the cataloging rules. Users are ill served by surrogates that are not rights aware. The Dublin Core element set has led the way for the library community by making rights management one of its 15 elements. Publishers are developing their own metadata standard, ONIX International Release 1.1., that also includes rights management elements. (Editeur) A new area is required in *AACR2* to address the, for all practical purposes, intrinsic needs of digital networked knowledge things for rights management information specific to them.

THE CHANGE FROM ASCERTAINING TO USING: A NEW PURPOSE FOR THE CATALOG

Changes to how we help others use packages of knowledge encounter *AACR2* in its heart of hearts: in the

role of the catalog record and the catalog itself as intermediaries between the book and its reader, between the resource and its user. The role of the catalog (and the records that populate it) is changing in two big ways. The first is the change from a finding aid to media delivery device to a virtual workspace. The second is the change from the premier research tool *in* the library to a valuable research tool in an Internet toolbox. The consequences of these two changes are far-reaching-they do change everything.

In a networked environment, the distances between the catalog record and the resource itself are annihilated. On the Web, the catalog record and the resource are hyperlinked together. They are not made one, but they are no longer independent and separated objects. The data and the metadata are physically *and* intellectually linked. A reader can use the surrogate to summon the resource itself. This sort of linkage has always been possible intellectually and imaginatively-quotations and citations are familiar examples of this. The Internet and hyper-linking make such intellectual or imaginative links real. Real in the sense that the links are physically present not just references and that the links can be used to make things happen-can bring the resource and the user together in a virtual workspace.

For networked resources, the catalog is not only a finding aid, a listing device. It is also multimedia delivery system. And it is more than that. A catalog on the Web is a portal to the Web. Like all such portals it is a door to a sub-set of the resources that populate the Internet. One library's portal may lead to a smaller or larger sub-set of resources than another, just as one Web search engine may index a smaller or larger sub-set than another engine. This is where the catalog is now, but it is not where it will stop. The development of the catalog will continue until it fulfils the promise of the fifth user need-use. The catalog must become a research tool that is integrated with the user's virtual workspace. The surrogates that populate our catalogs are no longer static and separate things. On this side of the digital divide, they are as dynamic and as linked-up as the resources they describe. This changes everything.

The Paris principles define the catalog as a tool for ascertaining whether or not some thing exists in a particular collection or collections. AACR1 is explicitly based on these principles. In *AACR2*, this basis is implied. Reduced to a single word, the purpose of the catalog is to *ascertain*. The *IFLA Functional Requirements for Bibliographic Records* defines the catalog in terms of meeting four user needs: to find, identify, select, and obtain. On this side of the digital divide, we must add to these four needs, a fifth: use.

For networked resources, display is insufficient. View, print, and save are only starting points. The catalog must deliver the networked resource to the user. Furthermore, it must do so in ways that enable the user to make use of the resource to meet the user primary needs. In an academic setting those primary needs are to teach, research, and publish. The catalog on the Web delivers networked resources to the user's virtual workspace, to the set of tools that enables the user to manipulate the resource-text, images, sounds, data, etc.-and put that content to their own uses.

Examples of such virtual workspaces that integrate data and metadata are now in use. NESSTAR, Networked Social Science Tools and Resources is one project developing such workspaces within the field of social science data archives. The NESSTAR project has developed sets of tools that allow

researchers to identify, locate, download, and use data from sites on the Web. The system is built upon DDI (Data Documentation Initiative)-compliant metadata. (NESSTAR Web Site) Another, from the field of digital art images, is Luna Imaging's Insight software at the Visual Resources Center at Yale University Library. Yale Library's implementation of Insight is a collaborative experiment in digital access to resources from the Yale University library and museum collections to support classroom teaching in the field of Art History. (Browser Insight) The Insight tool is builds upon VRA (Visual Resources Association) Core metadata. Tools such as these are the future of the catalog.

The new purposes of the catalog require a new conception of surrogates and catalogs, one that supports the linkage between the surrogate record and the resource. *AACR2* can no longer ignore the new bibliographic world that hyperlinks and networks are creating. Letting the MARC format carry the load, leaves the rules for cataloging less than Web aware. Adding a new area to *AACR2* (and to the ISBDs) for linking information is not the best approach. Since hyper-linking within a networked environment is a pervasive aspect of communication and publishing on this side of the digital divide, trying to keep linking in one area is counterproductive. The MARC format is already expanding the use of URLs in fields beyond the 856. In practice catalogers are far beyond even those extensions. The new purposes must be explicitly addressed in *AACR2* and linking must be supported throughout the rules.

In the analog age, the catalog has been the premier research tool in the library. In the digital networked age, it is a valuable research tool in an Internet toolbox. *AACR2* has grown up and flourished in the relatively homogenous confines of the library and its collections, purposes, services, traditions, and community of users. Other institutions did similar things but often did them differently, for different purposes, for different people, and in different places. Art museums and galleries, archives and natural history collections, indexing and abstracting services, research labs and projects are among the more obvious peers of libraries in collecting, organizing, and keeping sources of information to serve their users. In the digital networked era, libraries and their peer institutions are no longer so isolated from one another. The Internet has created opportunities for collaboration and even competition where few had existed before. The Web offers a heterogeneous world of resources to the researcher. Library catalogs are only one tool in this wider world.

The defining role of a library is that it is a collection or collections of selected materials. This is true in the analog era and in the digital era. Everything else we know about libraries and what they do relates back to this fundamental act (and fact) of selection. What is different in the digital era is that this role must be made explicitly clear to the user and not implied by the traditional limits of tangible things, books, buildings, campuses, etc. The shift from implicit landmarks to explicit signs is a generally applicable effect of the move to a digital networked environment. One implication for libraries is that catalogs cannot serve users well if they are conceived of as stand alone systems, as portals to one library's selection of Web resources. The catalog must be integrated with other resource discovery tools. For example, users of a catalog must also be able to turn their search into a broader Web search.

One aspect of the opening up has been the phenomenal interest in metadata. Another has been the development of myriad crosswalks to enable one metadata format to be translated into another. A third should be changes to *AACR2* that reflect this new world order. In homely terms, *AACR2* has been an only

child reared at home, and now *AACR2* has gone to nursery school and must learn to play well with others. Crosswalks are one way to play well with others, but they are exterior to the rules. Another way is to conceive of our rules as one way among many. We need to alter our rules and our principles so that we have the means to create records or surrogates that thrive within a rich, pluralistic world of dynamic and inter-linked resources and surrogates.

RECOMMENDATIONS

12 changes to make to *AACR2* to adapt it to a digital networked communications environment.

1. Change the purposes of the catalog by adding to the concepts in the Paris Principles those concepts of user needs expressed in the IFLA *FRBR*-find, identify, select, and obtain-and the fifth user need: *use*.
2. Change the concept of the catalog as a stand alone finding aid or listing device; explicitly state its ideal role as an tool designed to work well with other tools that use other metadata rules for their surrogates.
3. Change the orientation of display instructions from card production to online (hyper-linked) displays; change from editorial instructions to design guidance, include guidance for labeled and unlabeled displays, include explicit support for URLs and other reference linking techniques.
4. Use the concepts "work," "expression," "manifestation," and "item" as articulated in the IFLA report *Functional Requirements for Bibliographic Records* as a general framework within the rules. Concepts such as "edition," "impression," "original," or "copy" may continue to be highly useful for analog materials, but cannot be basic concepts of bibliographic control in an age of digital networked resources. Re-define the concept of a "reproduction" in an age of digital networked materials.
5. Thoroughly examine changing the arrangement of Part 1 of *AACR2* to follow an ISBD-like area order
6. Move instructions for cataloging particular types of materials out of the rules; collaborate with user communities to develop cataloging manuals (like *Bibliographic Description of Rare Books* and the *CONSER Cataloging Manual*) that are based on the rules.
7. Add new ISBD-like area for rights management information.
8. Add a new ISBD-like area for bibliographic relationships and reference linking.
9. Following the adoption of the proposed changes to chapter 12 (Serials), develop an ISBD-like area for the mode of issuance to include finite, serial, and integrating patterns of publication.
10. Eliminate *AACR2* chapter 9 (Electronic Resources); develop an ISBD-like area for the carrier aspects of all knowledge packages.
11. Further revise rule 0.24 so that the manifestation in hand or on screen remains the primary artifact being described; require that relations among manifestations or from manifestation to expression be articulated within or across records as needed.
12. Reconsider the role of transcription in descriptive cataloging. Since transcription is not an inherently suitable technique for describing dynamic or potentially dynamic resources, it may not be supportable as a primary means of creating identifiable surrogates.

REFERENCES

Association for Library Collections & Technical Services. Committee on Cataloging: Description and Access (1995). Guidelines for bibliographic description of reproductions. Chicago: American Library Association

Association for Library Collections & Technical Services. Committee on Cataloging: Description and Access (1994). Guidelines for bibliographic description of interactice multimedia. Chicago: American Library Association

Association for Library Collections & Technical Services. Committee on Cataloging: Description and Access. Task Force on an Appendix of Major and Minor Changes (2000). Report [Online] Available <http://www.ala.org/alcts/organization/ccs/ccda/tf-appx1.pdf> [Oct. 11, 2000]

Association for Library Collections & Technical Services. Machine-Readable Bibliographic Information Committee (MARBI). Proposal 93-4

Buckland, Michael (1991). "Information as thing." Journal of the American Society for Information Science. 42(5):351-360.

Borgman, Christine (2000). From Gutenberg to the global information infrastructure. Cambridge, Mass.: MIT Press, 2000.

Caplan, Priscilla and William Y. Arms (1999)."Reference linking for journal articles." D-lib magazine. 5(7/8) [Online] Available http://www.dlib.org/dlib/july99/caplan/07caplan.html [Oct. 11, 2000]

Committee on Intellectual Property Rights and the Emerging Information Infrastructure (2000). Digital Dilemma: Intellectual Property in the Information Age. Washington, D.C.: National Academy Press [Online] Available http://www.nap.edu/books/0309064996/html/ [Oct. 11, 2000]

Delsey, Tom (1998). The logical structure of the Anglo-American Cataloging Rules-Part I. [Online] Available <http://www.nlc-bnc.ca/jsc/> [Oct. 11, 2000]

Delsey, Tom (1999). The logical structure of the Anglo-American Cataloging Rules-Part II. [Online] Available <http://www.nlc-bnc.ca/jsc/> [Oct. 11, 2000]

Dublin Core Metadata Initiative (1998). Dublin Core metadata element set, version 1.1: reference description. [Online] Available <http://purl.org/dc/documents/rec-dces-199809.htm> [Oct. 11, 2000]

EDItEUR (2000) ONIX International Release 1.1 [Online] Available <http://www.editeur.org/onix.html> [Oct. 11, 2000]

Gladney, Henry M. (1999). "Digital dilemma: intellectual property." D-lib magazine. 5 (12) [Online] Available http://www.dlib.org/dlib/december99/gladney.html [Oct. 11, 2000]

Hirons, Jean and Crystal Graham (1998). 'Issues related to seriality', in The principles and future of AACR : proceedings of the International Conference on the Principles and Future Development of AACR, Toronto, Ontario, Canada, October 23-25, 1997. Ottawa : Chicago :Canadian Library Association; American Library Association.

Hirons, Jean (2000). Revising AACR to accommodate seriality: rule revision proposals. [Online] Available <http://www.nlc-bnc.ca/jsc/ch12.htm> [Oct. 11, 2000]

Hirons, Jean (1999). Revising *AACR2* to accommodate seriality : Report to the Joint Steering Committee for Revision of AACR, with the assistance of ... the CONSER AACR Review Task Force, April 1999. [Online] Available <http://www.nlc-bnc.ca/jsc/ser-rep0.html> [Oct. 11, 2000]

International Conference on Cataloguing Principles (1961 : Paris, France) (1971). Statement of principles: adopted at the International Conference on Cataloguing Principles, Paris, October 1961. London: IFLA Committee on Cataloging.

IFLA Study Group on the Functional Requirements for Bibliographic Records (1998). Functional requirements for bibliographic records final report. Munchen: Saur.

NESSTAR (Networked Social Science Tools and Resources) [Online] Available <http://www.nesstar.org/> [Oct. 11, 2000]

NetLibrary (2000) The MetaText digital textbook solution. [Online] Available <http://www.netlibrary.com/metatext.asp> [Oct. 11, 2000]

Taylor, Arlene G. (1999). The Organization of information. Englewood, Colo.: Libraries Unlimited.

Yale University Library. Imaging America Project. [Online] Available <http://www.library.yale.edu/art/Imaging_America.html> [Oct. 11, 2000]

Library of Congress
October 16, 2000

Comments: lcweb@loc.gov

Conference Home Page

What's new

Greetings from the Director for Cataloging

Topical discussion groups

NAS study and 2 articles from the LC staff *Gazette*

Conference program

Speakers, commentators, and papers

Conference sponsors

Conference discussion list

Logistical information for conference participants

Conference Organizing Team

Bicentennial Conference on Bibliographic Control for the New Millennium

Confronting the Challenges of Networked Resources and the Web

sponsored by the Library of Congress Cataloging Directorate

Matthew Beacom
Catalog Librarian for Network Information Resources
Sterling Memorial library
Yale University
130 Wall St.
New Haven, CT 06520

Crossing a Digital Divide: AACR2 and Unaddressed Problems of Networked Resources

About the presenter: Matthew Beacom has been a cataloger and a librarian for 10 years. He catalogs networked information resources for Yale University Library. Prior to his current position, Beacom cataloged books for the Beinecke Rare Book and Manuscript Library at Yale. He has been a member of ALCTS CC:DA (1996-2000) and a member of the PCC Standing Committee on Automation (1999-2001) and was once chair of the LITA/ALCTS Interest Group on Technical Sevices Workstations (1998). Beacom is currently a member of the ALTS CCS executive committee (2000-2002). Among the specific tasks he has worked on for professional committees are the ALCTS CC:DA Task Force on Harmonization of ISBD(ER) and AACR2 [1998-99] and the PCC SCA Task Group on Journals in Aggregator Databases.

Full text of paper is available

Summary:

The advent of the World Wide Web has initiated profound changes in how

BEST COPY AVAILABLE

Cataloging Directorate Home Page

Library of Congress Home Page

we use information. For librarians and those whom we serve, the most important changes may be in how new knowledge is created, how it is packaged, how it is published or disseminated, how its use is controlled, how it can be found, used, and saved for later use. In response to the Web and the cultural changes associated with it, librarians are thinking anew about how we enable people to have access to sources of information and knowledge. We are radically re-examining cataloging and catalogs.

In this paper, I address four problems or rather four changes in how we use knowledge that the library community must respond to. I address each change from the perspective of one who asks what the relation is between this change and AACR2, between this change and how libraries and librarians enable others to gain access to sources of information and knowledge. The four changes are:

1. The change in how knowledge is packaged,
2. The change in how knowledge is published and disseminated,
3. The change in how access to knowledge is controlled, and
4. The change in how we help others use knowledge as it is coming to be packaged, published, and restricted as networked resources.

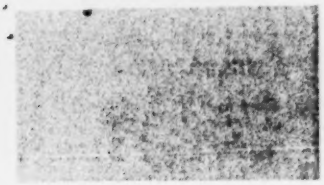
Changes to how knowledge is packaged as a networked resource encounter AACR2 most immediately in rule 0.24. These changes profoundly affect how we understand and resolve the relation between content and carrier and greatly multiply the scale of the multiple versions problem.

Changes to how packages of knowledge are published and disseminated encounter AACR2 most obviously in the publication area, but their impact is not limited to this area. Networked resources are packaged in new ways. E-journal aggregations are one. After 400 years, will journals continue to be the dominant delivery mechanism for articles? Networked resources have new qualities. They are egregiously updateable. For networked resources, updateable publications may become the dominant pattern.

Changes to how access to packages of knowledge is controlled encounter AACR2 in the note area. With networked resources, access restrictions are commonplace. With networked resources, the mix of universal and local information in bibliographic records is shifted toward local information. URLs for licensed materials demonstrate the importance of this shift.

Changes to how we help others use packages of knowledge encounter AACR2 in its heart of hearts, in the role of the catalog record and the catalog itself as intermediaries between the resource and the user, the book and its reader. In a networked environment, the distance--space and time--between the catalog record and the resource is annihilated. A catalog on the Web is a portal to the Web.

IFLA's *Functional Requirements for Bibliographic Records* defines four user needs: to find, to identify, to select, and to obtain. To these four we must add a fifth for networked resources: to use. For networked resources, the



catalog must deliver the resource to the user in ways that enable the user to make use of it to teach, do research, publish, etc. A catalog on the Web ideally delivers networked resources to the user's virtual workspace, a set of tools that enables the user to manipulate the resource--text, images, sounds, or data-- and put it to their own uses.